

FIGURE 1

CBH1-E1 Fusion Construct

T.reesei cbh1 core, linker (no CBD)

+

Acidothormus cellulolyticus endoglucanase 1 core (E1)

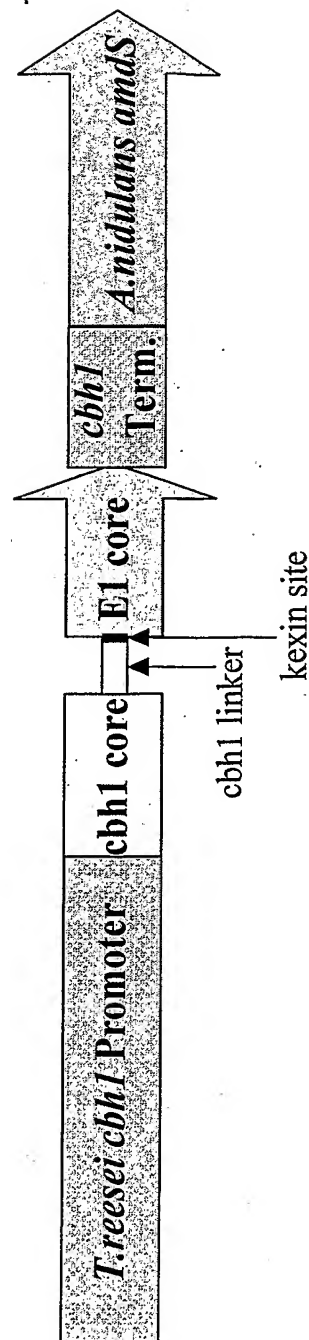


Figure 2

DNA sequence of *T.reesei cbh1* signal sequence+catalytic domain+linker (1570 bases)

ATGTATCGGAAGTTGGCCGTCATCTCGGCCTTCTTGGCCACAGCTCGTGCTCA
GTCGGCCTGCACTCTCCAATCGGAGACTCACCCGCCTCTGACATGGCAG
AAATGCTCGTCTGGTGGCACTTGCACCTCAACAGACAGGCTCCGTGGTCA
TCGACGCCAACTGGCGCTGGACTCACGCTACGAACAGCAGCACGAACTG
CTACGATGGCAACACTTGGAGCTCGACCCTATGTCCTGACAACGAGACC
TGCGCGAAGAAGTGTCTGTCTGGACGGTGCCGCCTACGCGTCCACGTACG
GAGTTACCACGAGCGGTAAACAGCCTCTCCATTGGCTTTGTACCCAGTC
TGCGCAGAAGAAGCTTGGCGCTCGCCTTTACCTTATGGCGAGCGACACG
ACCTACCAGGAATTCACCCTGCTTGGCAACGAGTTCTCTTTTCGATGTTGA
TGTTTCGCGAGCTGCCGTAAGTGACTTACCATGAACCCCTGACGTATCTTC
TTGTGGGCTCCCAGCTGACTGGCCAATTTAAGGTGCGGCTTGAACGGAG
CTCTCTACTTCGTGTCCATGGACGCGGATGGTGGCGTGAGCAAGTATCC
CACCAACACCGCTGGCGCCAAGTACGGCACGGGGTACTGTGACAGCCAG
TGTCCCCGCGATCTGAAGTTCATCAATGGCCAGGCCAACGTTGAGGGCT
GGGAGCCGTCATCCAACAACGCAAAACACGGGCATTGGAGGACACGGAA
GCTGCTGCTCTGAGATGGATATCTGGGAGGCCAACTCCATCTCCGAGGC
TCTTACCCCCCACCCTTGACGACTGTCGGCCAGGAGATCTGCGAGGGT
GATGGGTGCGGCGGAACCTTACTCCGATAACAGATATGGCGGGCACTTGCG
ATCCCGATGGCTGCGACTGGAACCCATACCGCCTGGGCAACACCAGCTT
CTACGGCCCTGGCTCAAGCTTTACCCTCGATACCACCAAGAAATTGACC
GTTGTACCCAGTTTCGAGACGTCGGGTGCCATCAACCGATACTATGTCC
AGAATGGCGTCACTTTCCAGCAGCCCAACGCCGAGCTTGGTAGTTACTC
TGGCAACGAGCTCAACGATGATTACTGCACAGCTGAGGAGGCAGAATTC
GGCGGATCCTCTTTCTCAGACAAGGGCGGCCTGACTCAGTTCAAGAAGG
CTACCTCTGGCGGCATGGTTCTGGTCATGAGTCTGTGGGATGATGTGAG
TTTGATGGACAAACATGCGCGTTGACAAAGAGTCAAGCAGCTGACTGAG
ATGTTACAGTACTACGCCAACATGCTGTGGCTGGACTCCACCTACCCGA
CAAACGAGACCTCCTCCACACCCGGTGCCGTGCGCGGAAGCTGCTCCAC
CAGCTCCGGTGTCCCTGCTCAGGTCGAATCTCAGTCTCCCAACGCCAAG
GTCACCTTCTCCAACATCAAGTTCGGACCCATTGGCAGCACCGGCAACC
CTAGCGGCGGCAACCCCTCCCGGCGGAAACCCGCCTGGCACCAACCACCCG
CCGCCCAGCCACTACCACTGGAAGCTCTCCCGGACCTACTAGT

Figure 3

Amino acid sequence of *T.reesei cbh1* signal sequence + catalytic domain + linker (480 amino acids)

MYRKLAVISAFLATARAQSACTLQSETHPPLTWQKCSSGGTCTQQTGSVVID
ANWRWTHATNSSTNCYDGNTWSSTLCPDNETCAKNCCLDGAAYASTYGVT
TSGNSLSIGFVTQSAQKNVGARLYLMASDTTYQEFTLLGNEFSFDVDVSQLP
CGLNGALYFVSMDADGGVSKYPTNTAGAKYGTGYCDSQCPRDLKFINGQA
NVEGWEPSSNNANTGIGGHGSCCSEMDIWEANSISEALTPHPCTTVGQEICE
GDGCGGTYSNRYGGTCDPDGCDWNPYRLGNTSFYGPSSFTLDTTKKLT
VVTQFETSGAINRYVQNGVTFQQPNAELGSYSGNELNDDYCTAEEAEFGG
SSFSDKGGLTQFKKATSGGMVLVMSLWDDYYANMLWLDSTYPTNETSSTP
GAVRGSCSTSSGVPAQVESQSPNAKVTFFSNIKFGPIGSTGNPSGGNPPGGNPPG
TTTTRRPATTGSSPGPTS

Figure 4

DNA sequence of *Acidothermus cellulolyticus* GH5A endoglucanase 1 catalytic domain
(1077 bases)

```
GCGGGCGGCGGCTATTGGCACACGAGCGGCCGGGAGATCCTGGACGCGAAC
AACGTGCCGGTACGGATCGCCGGGCATCAACTGGTTTGGGTTGCAAACCTGCA
ATTACGTCGTGCACGGTCTCTGGTCACGCGACTACCGCAGCATGCTCGACCA
GATAAAGTCGCTCGGCTACAACACAATCCGGCTGCCGTACTCTGACGACATT
CTCAAGCCGGGCACCATGCCGAACAGCATCAATTTTACCAGATGAATCAGG
ACCTGCAGGGTCTGACGTCCTTGAGGTCATGGACAAAATCGTCGCGTACGC
CGGTCAGATCGGCCTGCGCATCATTCTTGACCGCCACCGACCGGATTGCAGC
GGGCAGTCGGCGCTGTGGTACACGAGCAGCGTCTCGGAGGCTACGTGGATTT
CCGACCTGCAAGCGCTGGCGCAGCGCTACAAGGGAAACCCGACGGTCGTCG
GCTTTGACTTGCACAACGAGCCGCATGACCCGGCCTGCTGGGGCTGCGGCGA
TCCGAGCATCGACTGGCGATTGGCCGCCGAGCGGGCCGGAAACGCCGTGCTC
TCGGTGAATCCGAACCTGCTCATTTTCGTGCAAGGTGTGTCAGAGCTACAACG
GAGACTCCTACTGGTGGGGCGGCAACCTGCAAGGAGCCGGCCAGTACCCGGT
CGTGCTGAACGTGCCGAACCGCCTGGTGTACTCGGCGCACGACTACGCGACG
AGCGTCTACCCGCAGACGTGGTTCAGCGATCCGACCTTCCCAACAACATGC
CCGGCATCTGGAACAAGAACTGGGGATACCTCTTCAATCAGAACATTGCACC
GGTATGGCTGGGCGAATTCGGTACGACACTGCAATCCACGACCGACAGACG
TGGCTGAAGACGCTCGTCCAGTACCTACGGCCGACCGCGCAATACGGTGCGG
ACAGCTTCCAGTGGACCTTCTGGTCCTGGAACCCCGATTCCGGCGACACAGG
AGGAATTCTCAAGGATGACTGGCAGACGGTCGACACAGTAAAAGACGGCTAT
CTCGCGCCGATCAAGTCGTGATTTTCGATCCTGTGCGC
```

Figure 5

Amino acid sequence of *Acidothermus cellulolyticus* GH5A endoglucanase 1 catalytic domain (359 amino acids)

AGGGYWHTSGREILDANNVPVRIAGINWFGFETCNYYVHGLWSRDYRSMLDQI
KSLGYNTIRLPYSDDILKPGTMPNSINFYQMNQDLQGLTSLQVMDKIVAYAGQIG
LRIILDRHRPDCSGQSALWYTSSVSEATWISDLQALAQRYKGNPTVVGFDLHNEP
HDPACWGCGDPSIDWRLAAERAGNAVLSVNPNULLIFVEGVQSYNGDSYWWGG
NLQGAGQYPVVLNVPNRLVYSAHDYATSVYPQTWFS DPTFPNNMPGIWNKNW
GYLFNQNIAPVWLGEFGTTLQSTTDQTWLKT LVQYLRPTAQYGADSFQWTFWS
WNPDSGDTGGILKDDWQTVDTVKDGYLAPIKSSIFDPVG

FIGURE 6

DNA sequence of *Acidothermus cellulolyticus* GH48 cellulase catalytic domain (1914 bases)

AACGACCCGTACATCCAGCGGTTCTCACGATGTACAACAAGATTACAGACC
CAGCGAACGGCTACTTCAGCCCGCAGGGAATTCCCTACCACTCGGTAGAAAC
GCTCATCGTTGAGGCACCGGACTACGGGCACGAGACAACCTTCGGAGGCGTAC
AGCTTCTGGCTCTGGCTCGAAGCGACGTACGGCGCAGTGACCGGCAACTGGA
CGCCGTTCAACAACGCCTGGACGACGATGGAAACGTACATGATCCCGCAGCA
CGCGGACCAGCCGAACAACGCGTCGTACAACCCCAACAGCCCGGCGTCGTAC
GCTCCGGAAGAGCCGCTGCCCAGCATGTACCCGGTTGCCATCGACAGCAGCG
TGCCGGTTGGGCACGACCCGCTCGCCGCCGAATTGCAGTCGACGTACGGCAC
TCCGGACATTTACGGCATGCACTGGCTGGCCGACGTTGACAACATCTACGGA
TACGGCGACAGCCCCGGCGGTGGTTGCGAACTCGGTCCTTCCGCTAAGGGCG
TCTCCTACATCAACACATTCCAGCGCGGCTCGCAGGAGTCCGTCTGGGAGAC
GGTCACCCAGCCGACGTGCGACAACGGCAAGTACGGTGGGGCGCACGGCTA
CGTCGACCTGTTTCATCCAGGGTTCGACGCCGCCGACGTGGAAGTACACCGAT
GCCCCGGACGCCGACGCCCGTGCCGTCCAGGCTGCGTACTGGGCCTACACCT
GGGCATCGGCGCAGGGCAAGGCAAGCGCGATTGCCCCGACGATCGCCAAGG
CGGCCAAACCGGCGACTACCTGCGGTACTCGTCTTTGACAAGTACTTCAAG
CAGGTCGGCAACTGCTACCCGGCCAGCTCCTGCCCTGGAGCAACCGGACGCC
AGAGCGAGACCTACCTGATCGGCTGGTACTACGCCTGGGGCGGCTCAAGCCA
AGGCTGGGCCTGGCGCATTGGTGACGGCGCCGCGCACTTCGGCTACCAGAAT
CCGCTTGCCGCGTGGGCGATGTGCAACGTGACACCGCTCATTCCGCTCTCGCC
CACGGCAAAGAGCGACTGGGCGGCGAGCTTGACAGCGCCAGCTGGAGTTCTAC
CAGTGGTTGCAATCCGCGGAAGGAGCCATTGCGGGCGGCGCCACCAACAGCT
GGAACGGCAATTACGGGACCCCGCCGGCCGGAGACTCGACCTTCTACGGCAT
GGCGTACGACTGGGAGCCGGTCTACCACGACCCGCCGAGCAACAACCTGGTTC
GGCTTCCAGGCGTGGTCCATGGAACGGGTTGCCGAGTACTACTACGTCACCG
GCGACCCGAAGGCCAAGGCGCTGCTCGACAAGTGGGTGCGATGGGTGAAGC
CGAATGTCACCACCGGTGCCTCATGGTTCGATTCCGTCGAATTTGTCCTGGAGC
GGCCAACCGGATACCTGGAATCCGAGCAACCCAGGAACGAATGCCAACCTG
CACGTGACCATCACGTCGTCCGGGCAGGACGTCGGTGTGCGCGGGCGCTCG
CGAAGACACTCGAGTACTACGCGGCAAAATCCGGCGATACGGCCTCGCGCGA
CCTCGCGAAGGGATTGCTCGACTCCATGTGGAACAACGACCAGGACAGCCTC
GGTGTGAGCACACCGGAGACGCGGACCGACTACTCTCGGTTCACTCAGGTGT
ACGACCCGACGACTGGTGACGGCCTCTACATCCCGTCGGGTTGGACGGGGAC
CATGCCCAACGGTGACCAAATCAAGCCGGTGCGACCTTCCTGAGCATCCGG
TCCTGGTACACCAAGGATCCGCAGTGGTTCGAAGGTGCAGGCGTACCTCAACG
GCGGGCCTGCTCCGACGTTCAACTACCACCGGTTCTGGGCGGAGTCCGACTT
CGCGATGGCGAACGCCGATTTTGGCATGCTCTTCCCATCCGGG

FIGURE 7

Amino acid sequence of *Acidothermus cellulolyticus* 48 catalytic domain (638 amino acids)

NDPYIQRFLTMYNKHDPANGYFSPQGIPYHSVETLIVEAPDYGHETTSEAYSFW
LWLEATYGAVTGNWTPFNNAWTTMETYMIPQHADQPNNASYNPNSPASYAPEE
PLPSMYPVAIDSSVPVGHDPAAELQSTYGTPDIYGMHWLADVVDNIYGYGDSPG
GGCELGPSAKGVSYINTFQRGSQESVWETVTQPTCDNGKYGGAHGYVDLFIQGS
TPPWKYTDAPDADARAVQAAYWAYTWASAQGKASAIAPTIKAAKLGDYLR
YSLFDKYFKQVGNCYPASSCPGATGRQSETYLIGWYYAWGGSSQGWAWRIGD
GAAHFGYQNPLAAWAMSNVTPLIPLSPTAKSDWAASLQRQLEFYQWLQSAEGA
IAGGATNSWNGNYGTPPAGDSTFYGMAYDWEVYHDPPSNNWFGFQAWSMER
VAEYYYVTGDPKAKALLDKWVAWVKPNVTTGASWSIPSNLSWSGQPDTWNPS
NPGTNANLHVTITSSGQDVGVAAALAKTLEYA AAKSGDTASRDLAKGLLD SMW
NNDQDSLGVSTPETRTDYSRFTQVYDPTTGDGLYIPSGWTGTMPNGDQIKPGAT
FLSIRSWYTKDPQWSKVQAYLNGGPAPTFNYHRFWAESDFAMANADFGMLFPS
G

FIGURE 8A

DNA sequence of *Acidothermus cellulolyticus* GH74 catalytic domain

GCGACGACTCAGCCGTACACCTGGAGCAACGTGGCGATCGGGGGCGGCGGC
TTTGTGACGGGATCGTCTTCAATGAAGGTGCACCGGGAATTCTGTACGTGCG
GACGGACATCGGGGGGATGTATCGATGGGATGCCGCCAACGGGCGGTGGAT
CCCTCTTCTGGATTGGGTGGGATGGAACAATTGGGGGTACAACGGCGTCGTC
AGCATTGCGGCAGACCCGATCAATACTAACAAGGTATGGGCCGCCGTCGGAA
TGTACACCAACAGCTGGGACCCAAACGACGGAGCGATTCTCCGCTCGTCTGA
TCAGGGCGCAACGTGGCAAATAACGCCCTGCCGTTCAAGCTTGGCGGCAAC
ATGCCCCGGGCGTGGAATGGGCGAGCGGCTTGCGGTGGATCCAAACAATGACA
ACATTCTGTATTTTCGGCGCCCCGAGCGGCAAAGGGCTCTGGAGAAGCACAGA
TTCCGGCGCGACCTGGTCCCAGATGACGAACTTTCCGGACGTAGGCACGTAC
ATTGCAAATCCCCTGACACGACCGGCTATCAGAGCGATATTCAAGGCGTCG
TCTGGGTGCTTTTCGACAAGTCTTCGTCATCGCTCGGGCAAGCGAGTAAGACC
ATTTTTGTGGGCGTGCGGATCCCAATAATCCGGTCTTCTGGAGCAGAGACG
GCGGCGCGACGTGGCAGGCGGTGCCGGGTGCGCCGACCGGCTTCATCCCGCA
CAAGGGCGTCTTTGACCCGGTCAACCACGTGCTCTATATTGCCACCAGCAAT
ACGGGTGGTCCGTATGACGGGAGCTCCGGCGACGTCTGGAAATTCTCGGTGA
CCTCCGGGACATGGACGCGAATCAGCCCGGTACCTTCGACGGACACGGCCAA
CGACTACTTTGGTTACAGCGGCCTCACTATCGACCGCCAGCACCCGAACACG
ATAATGGTGGCAACCCAGATATCGTGGTGGCCGGACACCATAATCTTTTCGGA
GCACCGACGGCGGTGCGACGTGGACGCGGATCTGGGATTGGACGAGTTATCC
CAATCGAAGCTTGCGATATGTGCTTGACATTTTCGGCGGAGCCTTGGCTGACCT
TCGGCGTACAGCCGAATCCTCCCGTACCGAGTCCGAAGCTCGGCTGGATGGA
TGAAGCGATGGCAATCGATCCGTTCAACTCTGATCGGATGCTCTACGGAACA
GGCGCGACGTTGTACGCAACAAATGATCTCACGAAGTGGGACTCCGGCGGGCC
AGATTCATATCGCGCCGATGGTCAAAGGATTGGAGGAGACGGCGGTAAACG
ATCTCATCAGCCCGCCGTCTGGCGCCCCGCTCATCAGCGCTCTCGGAGACCTC
GGCGGCTTCACCCACGCCGACGTTACTGCCGTGCCATCGACGATCTTCACGTC

FIGURE 8B

ACCGGTGTTACGACCGGCACCAGCGTCGACTATGCGGAATTGAATCCGTCG
ATCATCGTTCGCGCTGGAAGTTTCGATCCATCGAGCCAACCGAACGACAGGC
ACGTCGCGTTCTCGACAGACGGCGGCAAGAACTGGTTCCAAGGCAGCGAACC
TGGCGGGGTGACGACGGGCGGCACCGTCGCCGCATCGGCCGACGGCTCTCGT
TTCGTCTGGGCTCCCGGCGATCCCGGTCAGCCTGTGGTGTACGCAGTCGGATT
TGGCAACTCCTGGGCTGCTTCGCAAGGTGTTCCCGCCAATGCCCAGATCCGCT
CAGACCGGGTGAATCCAAAGACTTTCTATGCCCTATCCAATGGAACCTTCTAT
CGAAGCACGGACGGCGGCGTGACATTCCAACCGGTCGCGGCCGGTCTTCCGA
GCAGCGGTGCCGTCGGTGTTCATGTTCCACGCGGTGCCTGGAAAAGAAGGCCGA
TCTGTGGCTCGCTGCATCGAGCGGGCTTTACCACTCAACCAATGGCGGCAGC
AGTTGGTCTGCAATCACCGGCGTATCCTCCGCGGTGAACGTGGGATTTGGTA
AGTCTGCGCCCGGGTCGTCATACCCAGCCGTCTTTGTCGTCGGCACGATCGGA
GGCGTTACGGGGGCGTACCGCTCCGACGACGGTGGGACGACCTGGGTACGG
ATCAATGATGACCAGCACCAATACGGAAATTGGGGACAAGCAATCACCGGTG
ACCCGCGAATTTACGGGCGGGTGTACATAGGCACGAACGGCCGTGGAATTGT
CTACGGGGACATTGGTGGTGCGCCGTCCGGATCG

FIGURE 9

Amino acid sequence of *Acidothermus cellulolyticus* 74 catalytic domain (741 amino acids)

ATTQPYTWSNVAIGGGGFVDGIVFNEGAPGILYVRTDIGGMYRWDAANGRWIPL
LDWVGWNNWGYNGVVSIADPINTNKVWAAVGMYTNSWDPNDGAILRSSDQ
GATWQITPLPFKLGGNMPGRGMGERLAVDPNNDNILYFGAPSGKGLWRSTDG
ATWSQMTNFPDVGTYIANPTDTTGYQSDIQGVVWVAFDKSSSSLGQASKTIFVG
VADPNNPVFWSRDGGATWQAVPGAPTGFIPHKGVFDPVNHVLYIATSNTGGPY
DGSSGDVWKFSVTSGTWTRISPVPSTDTANDYFGYSGLTIDRQHPNTIMVATQIS
WWPDTIIFRSTDGGATWTRIWDWTSYPNRSRLRYVLDISAEPWLTFGVQPNPPVPS
PKLGWMDEAMAIDPFNSDRMLYGTGATLYATNDLTKWDSGGQIHIAPMVKGLE
ETA VNDLISPPSGAPLISALGDLGGFTHADVTAVPSTIFTSPVFTTGTSVDY AELNP
SIIVRAGSFDPSQPNDRHVAFSTDGGKNWFQGSEPGGVTTGGTVAASADGSRFV
WAPGDPGQPVVYAVGFGNSWAASQGV PANAQIRSDRVNPKTFYALSNGTFYRS
TDGGVTFQPVAAGLPSSGAVGVMFHAVPGKEGDLWLAASSGLYHSTNGGSSWS
AITGVSSAVNVGFGKSAPGSSYPV FVVGTTGGVTGAYRSDDGGTTWVRINDDQ
HQYGNW GQAITGDPRIYGRVYIGTNGRGIVYGDIGGAPSGS

Figure 10

DNA sequence of *Thermobifida fusca* E3 (TfE3) cellulase including the cellulose binding domain - linker - catalytic domain but lacking a TfE3 signal sequence. (1677 bases)

```
GCCGGCTGCTCGGTGGACTACACGGTCAACTCCTGGGGTACCGGGTTCACCG
CCAACGTACCATCACCAACCTCGGCAGTGCGATCAACGGCTGGACCCTGGA
GTGGGACTTCCCCGGCAACCAGCAGGTGACCAACCTGTGGAACGGGACCTAC
ACCCAGTCCGGGCAGCACGTGTCGGTCAGCAACGCCCCGTACAACGCCTCCA
TCCCCGGCCAACGGAACGGTTGAGTTCGGGTTCAACGGCTCCTACTCGGGCAG
CAACGACATCCCCTCCTCCTTCAAGCTGAACGGGGTTACCTGCGACGGCTCG
GACGACCCCGACCCCGAGCCCGAGCCCTCCCCCAGCCCTTCCCCCAGCCCCA
CAGACCCGGATGAGCCGGGCGGCCCGACCAACCCGCCACCAACCCCGGCG
AGAAGGTGACAAACCCGTTTCGAGGGCGCCAAGCTGTACGTGAACCCGGTCTG
GTCGGCCAAGGCCGCGCTGAGCCGGGCGGTTCCGCGGTTCGCCAACGAGTCC
ACCGCTGTCTGGCTGGACCGTATCGGCGCCATCGAGGGCAACGACAGCCCGA
CCACCGGCTCCATGGGTCTGCGCGACCACTGGAGGAGGCCGTCCGCCAGTC
CGGTGGCGACCCGCTGACCATCCAGGTCGTCTATCTACAACCTGCCCGGCCGC
GACTGCGCCGCGCTGGCCTCCAACGGTGAGCTGGGTCCCGATGAACTCGACC
GCTACAAGAGCGAGTACATCGACCCGATCGCCGACATCATGTGGGACTTCGC
AGACTACGAGAACCTGCGGATCGTCGCCATCATCGAGATCGACTCCCTGCCC
AACCTCGTCACCAACGTGGGCGGGAACGGCGGCACCGAGCTCTGCGCCTACA
TGAAGCAGAACGGCGGCTACGTCAACGGTGTCGGCTACGCCCTCCGCAAGCT
GGGCGAGATCCCGAACGTCTACAACCTACATCGACGCCGCCACACGGCTGG
ATCGGCTGGGACTCCAACCTTCGGCCCCCTCGGTGGACATCTTCTACGAGGCCG
CCAACGCCTCCGGCTCCACCGTGGACTACGTGCACGGCTTCATCTCCAACAC
GGCCAACTACTCGGCCACTGTGGAGCCGTACCTGGACGTCAACGGCACCCGTT
AACGGCCAGCTCATCCGCCAGTCCAAGTGGGTTGACTGGAACCAGTACGTCTG
ACGAGCTCTCCTTCGTCCAGGACCTGCGTCAGGCCCTGATCGCCAAGGGCTTC
CGGTCCGACATCGGTATGCTCATCGACACCTCCCGCAACGGCTGGGGTGGCC
CGAACCGTCCGACCGGACCGAGCTCCTCCACCGACCTCAACACCTACGTTGA
CGAGAGCCGTATCGACCGCCGTATCCACCCCGGTAACCTGGTGCAACCAGGCC
GGTGCGGGCCTCGGCGAGCGGCCACGGTCAACCCGGCTCCCGGTGTTGACG
CCTACGTCTGGGTGAAGCCCCCGGGTGAGTCCGACGGCGCCAGCGAGGAGAT
CCCGAACGACGAGGGCAAGGGCTTCGACCGCATGTGCGACCCGACCTACCAG
GGCAACGCCCGCAACGGCAACAACCCCTCGGGTGCGCTGCCCAACGCCCCCA
TCTCCGGCCACTGGTTCTCTGCCAGTTCCGCGAGCTGCTGGCCAACGCCTAC
CCGCCTCTGTAA
```

Figure 11

Amino acid sequence of the *Thermobifida fusca* E3 - cellulase including the cellulose binding domain - linker - catalytic domain but lacking the Tfe3 signal sequence. (558 amino acids)

AGCSVDYTVNSWGTGFTANVTITNLGSAINGWTLEWDFPGNQQTNLWNGTYT
QSGQHVSVSNAPYNASIPANGTVEFGFNGSYSGSNDIPSSFKLNGVTCDSDDPD
PEPSPSPSPSPSTDPDEPGGPTNPPTNPGEKVDNPFEGAKLYVNPVWSAKAAAEP
GGSANESTAVWLDRIGAIEGNDSPPTGSMGLRDHLEEAVRQSGGDPLTIQVVI
YNLPGRDCAALASNGELGPDELDRYKSEYIDPIADIMWDFADYENLRIVAIEIDS
LPNLVTNVGGNGGTELCAVMKQNGGYVNGVGYALRKLGEIPNVYNYIDAHH
GWIGWDSNFGPSVDIFYEAAANASGSTVDYVHGFISNTANYSATVEPYLDVNGTV
NGQLIRQSKWVDWNQYVDELSFVQDLRQALIAKGFRSDIGMLIDTSRNGWGGP
NRPTGPSSSTDLNTYVDESRIDRRIHPGNWCNQAGAGLGERPTVNPAPGVDAVV
WVKPPGESDGASEEIPNDEGKGFDRCMDPTYQGNARNGNNPSGALPNAPISGH
WFSAQFRELLANAYPPL

Figure 12

DNA sequence of *Thermobifida fusca* E5 (TfE5) endoglucanase including the cellulose binding domain - linker and catalytic domain but lacking a TfE5 signal sequence. (1293 bases)

```
GCCGGTCTCACCGCCACAGTCACCAAAGAATCCTCGTGGGACAACGGCTACT
CCGCGTCCGTACCGTCCGCAACGACACCTCGAGCACCGTCTCCCAGTGGGA
GGTCGTCCTCACCTGCCCGGCGGCACTACAGTGGCCCAGGTGTGGAACGCC
CAGCACACCAGCAGCGGCAACTCCACACCTTCACCGGGGTTTCCTGGAACA
GCACCATCCCCGCCGGAGGCACCGCCTCTTCCGGCTTCATCGCTTCCGGCAGC
GGCGAACCCACCCACTGCACCATCAACGGCGCCCCCTGCGACGAAGGCTCCG
AGCCGGGCGGCCCCGGCGGTCCCGGAACCCCTCCCCGACCCCGGCACGCA
GCCCCGCACCGGCACCCCGGTTCGAGCGGTACGGCAAAGTCCAGGTCTGCGGC
ACCCAGCTCTGCGACGAGCAGGCAACCCGGTCCAAGTGCAGCGGCATGAGCA
CCCACGGCATCCAGTGGTTCGACCACTGCCTGACCGACAGCTCGCTGGACGC
CCTGGCCTACGACTGGAAGGCCGACATCATCCGCCTGTCCATGTACATCCAG
GAAGACGGCTACGAGACCAACCCGCGCGGCTTCACCGACCGGATGCACCAG
CTCATCGACATGGCCACGGCGCGCGGCCTGTACGTGATCGTGGACTGGCACA
TCCTCACCCCGGGCGATCCCCACTACAACCTGGACCGGGCCAAGACCTTCTTC
GCGGAAATCGCCCAGCGCCACGCCAGCAAGACCAACGTGCTCTACGAGATCG
CCAACGAACCCAACGGAGTGAGCTGGGCCTCCATCAAGAGCTACGCCGAAG
AGGTCATCCCGGTGATCCGCCAGCGCGACCCCGACTCGGTGATCATCGTGGG
CACCCGCGGCTGGTCGTCGCTCGGCGTCTCCGAAGGCTCCGGCCCCGCGGAG
ATCGCGGCCAACCCGGTCAACGCCTCCAACATCATGTACGCCTTCCACTTCTA
CGCGGCCTCGCACCGCGACAACTACCTCAACGCGCTGCGTGAGGCCTCCGAG
CTGTTCCCGGTCTTCGTCACCGAGTTCGGCACCGAGACCTACACCGGTGACG
GCGCCAACGACTTCCAGATGGCCGACCGCTACATCGACCTGATGGCGGAACG
GAAGATCGGGTGGACCAAGTGGAATACTCGGACGACTTCCGTTCCGGCGCG
GTCTTCCAGCCGGGCACCTGCGCGTCCGGCGGGCCCGTGGAGCGGTTTCGTCGC
TGAAGGCGTCCGGACAGTGGGTGCGGAGCAAGCTCCAGTCCTGA
```

Figure 13

Amino acid sequence of the *Thermobifida fusca* E5 –cellulase including the cellulose binding domain - linker - catalytic domain but lacking a TfE5 signal sequence. (430 amino acids)

AGLTATVTKESSWDNGYSASVTVRNDTSSTVSQWEVVLTLPGGTTVAQVWNAQ
HTSSGNSHTFTGVSWNSTIPPGGTASSGFIASGSGEPHCTINGAPCDEGSEPGGP
GGPGTPSPDPGTQPGTGTPVERYGKVQVCGTQLCDEHGNPVQLRGMSTHGIQW
FDHCLTDSSLDALAYDWKADIIRLSMYIQEDGYETNPRGFTDRMHQLIDMATAR
GLYVIVDWHILTPGDPHYNLDRAKTFFAEIAQRHASKTNVLYEIANEPNGVSWA
SIKSYAEEVIPVIRQRPDSVIIVGTRGWSSLGVSESGPAEIAANPVNASNIMYAF
HFYAASHRDNYLNALREASELFPVFVTEFGTETYTGDGANDFQMADRYIDLMA
ERKIGWTKWNYSDDFRSGAVFQPGTCASGGPWSSSLKASGQWVRSKLQS

Figure 14

DNA sequence of CBH1-E1 fusion (2656 bases)

T.reesei cbh1 signal sequence+catalytic domain+linker+added amino acidsSKR+*Acidothermus cellulolyticus* GH5A catalytic domain

ATGTATCGGAAGTTGGCCGTCATCTCGGCCTTCTTGGCCACAGCTCGTGCTCA
GTCGGCCTGCACTCTCCAATCGGAGACTCACCCGCCTCTGACATGGCAGAAA
TGCTCGTCTGGTGGCACTTGCACTCAACAGACAGGCTCCGTGGTCATCGACG
CCAACTGGCGCTGGACTCACGCTACGAACAGCAGCACGAAGTCTACGATGG
CAACACTTGGAGCTCGACCCTATGTCCTGACAACGAGACCTGCGCGAAGAAC
TGCTGTCTGGACGGTGCCGCCTACGCGTCCACGTACGGAGTTACCACGAGCG
GTAACAGCCTCTCCATTGGCTTTGTACCCAGTCTGCGCAGAAGAACGTTGGC
GCTCGCCTTTACCTTATGGCGAGCGACACGACCTACCAGGAATTCACCCTGCT
TGGCAACGAGTTCTCTTTTCGATGTTGATGTTTCGCAGCTGCCGTAAGTGACTT
ACCATGAACCCCTGACGTATCTTCTTGTGGGCTCCCAGCTGACTGGCCAATTT
AAGGTGCGGCTTGAACGGAGCTCTCTACTTCGTGTCCATGGACGCGGATGGT
GGCGTGAGCAAGTATCCCAACACACCGCTGGCGCCAAGTACGGCACGGGGT
ACTGTGACAGCCAGTGTCCCCGCGATCTGAAGTTCATCAATGGCCAGGCCAA
CGTTGAGGGCTGGGAGCCGTCATCCAACAACGCAAACACGGGCATTGGAGG
ACACGGAAGCTGCTGCTCTGAGATGGATATCTGGGAGGCCAACTCCATCTCC
GAGGCTCTTACCCCCCACCCTTGACGACTGTGCGGCCAGGAGATCTGCGAGG
GTGATGGGTGCGGCGGAACCTTACTCCGATAACAGATATGGCGGCACTTGCGA
TCCCGATGGCTGCGACTGGAACCCATACCGCCTGGGCAACACCAGCTTCTAC
GGCCCTGGCTCAAGCTTTACCCTCGATACCACCAAGAAATTGACCGTTGTAC
CCAGTTCGAGACGTCGGGTGCCATCAACCGATACTATGTCCAGAATGGCGTC
ACTTTCCAGCAGCCCAACGCCGAGCTTGGTAGTTACTCTGGCAACGAGCTCA
ACGATGATTACTGCACAGCTGAGGAGGCAGAATTCGGCGGATCCTCTTTCTC
AGACAAGGGCGGCCTGACTCAGTTCAAGAAGGCTACCTCTGGCGGCATGGTT
CTGGTCATGAGTCTGTGGGATGATGTGAGTTTGATGGACAAACATGCGCGTT
GACAAAGAGTCAAGCAGCTGACTGAGATGTTACAGTACTACGCCAACATGCT
GTGGCTGGACTCCACCTACCCGACAAACGAGACCTCCTCCACACCCGGTGCC
GTGCGCGGAAGCTGCTCCACCAGCTCCGGTGTCCCTGCTCAGGTGCAATCTC
AGTCTCCCAACGCCAAGGTACCTTCTCCAACATCAAGTTCGGACCCATTGGC
AGCACCGGCAACCCTAGCGGCGGCAACCCTCCCGGCGGAAACCCGCCTGGCA
CCACCACCACCCGCCGCCAGCCACTACCACTGGAAGCTCTCCCGGACCTAC
TAGTAAGCGGGCGGGCGGCGGCTATTGGCACACGAGCGGCCGGGAGATCCT
GGACGCGAACAACGTGCCGTACGGATCGCCGGCATCAACTGGTTTGGGTTT
GAAACCTGCAATTACGTCTGTCACGGTCTCTGGTCACGCGACTACCGCAGCA
TGCTCGACCAGATAAAGTCGCTCGGCTACAACACAATCCGGCTGCCGTACTC
TGACGACATTCTCAAGCCGGGCACCATGCCGAACAGCATCAATTTTACCAG
ATGAATCAGGACCTGCAGGGTCTGACGTCCTTGACAGGTTCATGGACAAAATCG
TCGCGTACGCCGGTCAGATCGGCCTGCGCATCATTCTTGACCGCCACCGACC
GGATTGCAGCGGGCAGTCGGCGCTGTGGTACACGAGCAGCGTCTCGGAGGCT
ACGTGGATTTCGACCTGCAAGCGCTGGCGCAGCGCTACAAGGGAAACCCGA

CGGTCGTCGGCTTTGACTTGCACAACGAGCCGCATGACCCGGCCTGCTGGGG
CTGCGGCGATCCGAGCATCGACTGGCGATTGGCCGCCGAGCGGGCCGGAAC
GCCGTGCTCTCGGTGAATCCGAACCTGCTCATTTTCGTCGAAGGTGTGCAGAG
CTACAACGGAGACTCCTACTGGTGGGGCGGCAACCTGCAAGGAGCCGGCCA
GTACCCGGTCGTGCTGAACGTGCCGAACCGCCTGGTGTACTCGGCGCACGAC
TACGCGACGAGCGTCTACCCGCAGACGTGGTTCAGCGATCCGACCTTCCCA
ACAACATGCCCCGGCATCTGGAACAAGAACTGGGGATACCTCTTCAATCAGAA
CATTGCACCGGTATGGCTGGGCGAATTCGGTACGACACTGCAATCCACGACC
GACCAGACGTGGCTGAAGACGCTCGTCCAGTACCTACGGCCGACCGCGCAAT
ACGGTGCGGACAGCTTCCAGTGGACCTTCTGGTCCTGGAACCCCGATTCCGG
CGACACAGGAGGAATTCTCAAGGATGACTGGCAGACGGTCGACACAGTAAA
AGACGGCTATCTCGCGCCGATCAAGTCGTCGATTTTCGATCCTGTCGGCTAA

Figure 15

Amino acid sequence of CBH1-E1 fusion (841 amino acids)

T.reesei cbh1 signal sequence+catalytic domain+linker+added amino acids

SKR+*Acidothermus cellulolyticus* GH5A catalytic domain

MYRKLAVISAFLATARAQSACTLQSETHPPLTWQKCSSGGTCTQQTGSVVIDAN
WRWTHATNSSTNCYDGNTWSSTLCPDNETCAKNCCLDGAAYASTYGVTTSGNS
LSIGFVTQSAQKNVGARLYLMASDTTYQEFTLLGNEFSFDVDVSQLPCGLNGAL
YFVSMDADGGVSKYPTNTAGAKYGTGYCDSQCPRDLKFINGQANVEGWEPSSN
NANTGIGGHGSCCSEMDIWEANSISEALTPHPCTTVGQEICEGDGCGGTYSNRY
GGTCDPDGCDWNPYRLGNTSFYGPSSFTLDTTKKLTVVTQFETSGAINRYYVQ
NGVTFQQPNAELGSYSGNELNDDYCTAEAEFGSSFSKGGTLQFKKATSGGM
VLVMSLWDDYYANMLWLDSTYPTNETSSTPGAVRGSCSTSSGVPAQVESQSPN
AKVTFSNIKFGPIGSTGNPSGGNPPGGNPPGTTTTTRRPATTTGSSPGPTSKRAGGG
YWHTSGREILDANNVPVRIAGINWFGFETCNVYVHGLWSRDYRSMLDQIKSLGY
NTIRLPYSDDILKPGTMPNSINFYQMNQDLQGLTSLQVMDKIVAYAGQIGLRILD
RHRPDCSGQSALWYTSSVSEATWISDLQALAQRYKGNPTVVGFDLHNEPHDPAC
WGCGDPSIDWRLAAERAGNAVLSVNPPLLIFVEGVQSYNGDSYWWGGNLQGA
GQYPVVLNVPNRLVYSAHDYATSVYPQTWFSPTFPNNMPGIWNKNWGYLFN
QNIAPVWLGEFGTTLQSTTDQTLWLKTLVQYLRPTAQYGADSFQWTFWSWNPDS
GDTGGILKDDWQTVDTVKDGYLAPIKSSIFDPVG

Figure 16

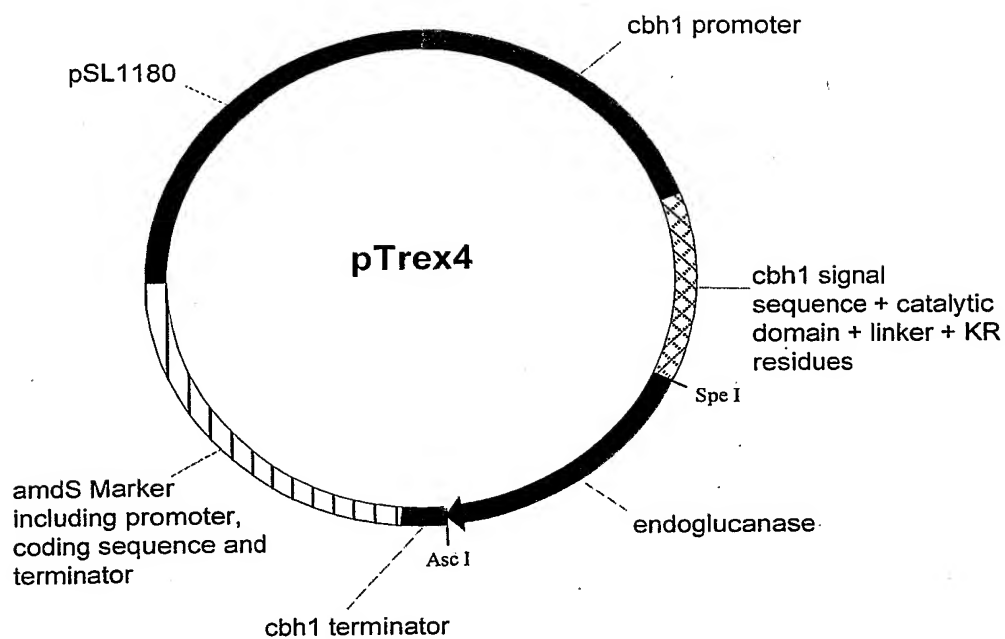


Figure 17

DNA sequence of pTrex4 (10239 bases)

AAGCTTAACTAGTACTTCTCGAGCTCTGTACATGTCCGGTCGCGACGTACGCG
TATCGATGGCGCCAGCTGCAGGCGGCCGCCTGCAGCCACTTGCAGTCCCGTG
GAATTCTCAGCGTGAATGTAGGCCTTTTGTAGGGTAGGAATTGTCACTCAAGC
ACCCCCAACCTCCATTACGCCTCCCCCATAGAGTTCCCAATCAGTGAGTCATG
GCACTGTTCTCAAATAGATTGGGGAGAAGTTGACTTCCGCCCAGAGCTGAAG
GTCGCACAACCGCATGATATAGGGTCGGCAACGGCAAAAAAGCACGTGGCT
CACCGAAAAGCAAGATGTTTTCGATCTAACATCCAGGAACCTGGATACATCC
ATCATCACGCACGACCACTTTGATCTGCTGGTAAACTCGTATTCGCCCTAAAC
CGAAGTGACGTGGTAAATCTACACGTGGGCCCCTTTCGGTATACTGCGTGTGT
CTTCTCTAGGTGCCATTCTTTTCCCTTCTCTAGTGTTGAATTGTTTGTGTTGG
AGTCCGAGCTGTAACCTCTGAATCTCTGGAGAATGGTGGACTAACGACT
ACCGTGCACCTGCATCATGTATATAATAGTGATCCTGAGAAGGGGGGTTTGG
AGCAATGTGGGACTTTGATGGTCATCAAACAAAGAACGAAGACGCCTCTTTT
GCAAAGTTTTGTTTCGGCTACGGTGAAGAACTGGATACTTGTTGTGTCTTCTG
TGTATTTTTGTGGCAACAAGAGGCCAGAGACAATCTATTCAAACACCAAGCT
TGCTCTTTTGAGCTACAAGAACCTGTGGGGTATATATCTAGAGTTGTGAAGTC
GGTAATCCCGCTGTATAGTAATACGAGTCGCATCTAAATACTCCGAAGCTGCT
GCGAACCCGGAGAATCGAGATGTGCTGGAAAGCTTCTAGCGAGCGGCTAAAT
TAGCATGAAAGGCTATGAGAAATTCTGGAGACGGCTTGTTGAATCATGGCGT
TCCATTCTTCGACAAGCAAAGCGTTCCGTCGCAGTAGCAGGCACTCATTCCCG
AAAAAACTCGGAGATTCCTAAGTAGCGATGGAACCGGAATAATATAATAGGC
AATACATTGAGTTGCCTCGACGGTTGCAATGCAGGGGTACTGAGCTTGGACA
TAACTGTTCCGTACCCACCTCTTCTCAACCTTTGGCGTTTCCCTGATTACGCG
TACCCGTACAAGTCGTAATCACTATTAACCCAGACTGACCGGACGTGTTTTGC
CCTTCATTTGGAGAAATAATGTCATTGCGATGTGTAATTTGCCTGCTTGACCG
ACTGGGGCTGTTTCAAGCCCGAATGTAGGATTGTTATCCGAACCTCTGCTCGTA
GAGGCATGTTGTGAATCTGTGTCGGGCAGGACACGCCTCGAAGGTTACGGC

AAGGGAAACCAACCGATAGCAGTGTCTAGTAGCAACCTGTAAAGCCGCAATGC
AGCATCACTGGAAAATACAAACCAATGGCTAAAAGTACATAAGTTAATGCCT
AAAGAAGTCATATAACCAGCGGCTAATAATTGTACAATCAAGTGGCTAAACGT
ACCGTAATTTGCCAACGGCTTGTGGGGTTGCAGAAGCAACGGCAAAGCCCCA
CTTCCCCACGTTTGTTCCTTCACTCAGTCCAATCTCAGCTGGTGATCCCCCAAT
TGGGTCGCTTGTTCGCGGTGAAGTGAAAGAAGACAGAGGTAAGAATGTC
TGACTCGGAGCGTTTTGCATACAACCAAGGGCAGTGATGGAAGACAGTGAAA
TGTTGACATTCAAGGAGTATTTAGCCAGGGATGCTTGAGTGTATCGTGTAAG
GAGGTTTGTCTGCCGATACGACGAATACTGTATAGTCACTTCTGATGAAGTGG
TCCATATTGAAATGTAAGTCGGCACTGAACAGGCAAAAGATTGAGTTGAAAC
TGCCTAAGATCTCGGGCCCTCGGGCCTTCGGCCTTTGGGTGTACATGTTTGTG
CTCCGGGCAAATGCAAAGTGTGGTAGGATCGAACACACTGCTGCCTTTACCA
AGCAGCTGAGGGTATGTGATAGGCAAATGTTTCAGGGGCCACTGCATGGTTTC
GAATAGAAAGAGAAGCTTAGCCAAGAACAATAGCCGATAAAGATAGCCTCA
TTAAACGGAATGAGCTAGTAGGCAAAGTCAGCGAATGTGTATATATAAAGGT
TCGAGGTCCGTGCCTCCCTCATGCTCTCCCCATCTACTCATCAACTCAGATCC
TCCAGGAGACTTGTACACCATCTTTTGAGGCACAGAAACCCAATAGTCAACC
GCGGACTGCGCATCATGTATCGGAAGTTGGCCGTCATCTCGGCCTTCTTGGCC
ACAGCTCGTGCTCAGTCGGCCTGCACTCTCCAATCGGAGACTCACCCGCCTCT
GACATGGCAGAAATGCTCGTCTGGTGGCACTTGCACTCAACAGACAGGCTCC
GTGGTCATCGACGCCAACTGGCGCTGGACTCACGCTACGAACAGCAGCACGA
ACTGCTACGATGGCAACACTTGGAGCTCGACCCTATGTCCTGACAACGAGAC
CTGCGCGAAGAACTGCTGTCTGGACGGTGCCGCCTACGCGTCCACGTACGGA
GTTACCACGAGCGGTAACAGCCTCTCCATTGGCTTTGTCACCCAGTCTGCGCA
GAAGAACGTTGGCGCTCGCCTTTACCTTATGGCGAGCGACACGACCTACCAG
GAATTCACCCTGCTTGGCAACGAGTTCTCTTTCGATGTTGATGTTTCGCAGCT
GCCGTAAGTGACTTACCATGAACCCCTGACGTATCTTCTTGTGGGCTCCCAGC
TGACTGGCCAATTTAAGGTGCGGCTTGAACGGAGCTCTCTACTTCGTGTCCAT
GGACGCGGATGGTGGCGTGAGCAAGTATCCCACCAACACCGCTGGCGCCAA
GTACGGCACGGGGTACTGTGACAGCCAGTGTCCCCGCGATCTGAAGTTCATC
AATGGCCAGGCCAACGTTGAGGGCTGGGAGCCGTCATCCAACAACGCAAAC

ACGGGCATTGGAGGACACGGAAGCTGCTGCTCTGAGATGGATATCTGGGAGG
CCAACTCCATCTCCGAGGCTCTTACCCCCACCCTTGACGACTGTCGGCCAG
GAGATCTGCGAGGGTGATGGGTGCGGCGGAACCTTACTCCGATAACAGATATG
GCGGCACTTGCGATCCCGATGGCTGCGACTGGAACCCATACCGCCTGGGCAA
CACCAGCTTCTACGGCCCTGGCTCAAGCTTTACCCTCGATACCACCAAGAAAT
TGACCGTTGTACCCAGTTCGAGACGTCGGGTGCCATCAACCGATACTATGTC
CAGAATGGCGTCACTTTCCAGCAGCCCAACGCCGAGCTTGGTAGTTACTCTG
GCAACGAGCTCAACGATGATTACTGCACAGCTGAGGAGGCAGAATTCGGCGG
ATCCTCTTTCTCAGACAAGGGCGGCCTGACTCAGTTCAAGAAGGCTACCTCTG
GCGGCATGGTTCTGGTCATGAGTCTGTGGGATGATGTGAGTTTGATGGACAA
ACATGCGCGTTGACAAAGAGTCAAGCAGCTGACTGAGATGTTACAGTACTAC
GCCAACATGCTGTGGCTGGACTCCACCTACCCGACAAACGAGACCTCCTCCA
CACCCGGTGCCGTGCGCGGAAGCTGCTCCACCAGCTCCGGTGTCCCTGCTCA
GGTCGAATCTCAGTCTCCCAACGCCAAGGTCACCTTCTCCAACATCAAGTTCG
GACCCATTGGCAGCACCGGCAACCCTAGCGGCGGCAACCCTCCCGGCGGAAA
CCCGCCTGGCACCAACCACCAACCCGCCGCCAGCCACTACCACTGGAAGCTCT
CCCGGACCTACTAGTAAGCGGATAAGGCGCGCCGCGCGCCAGCTCCGTGCGA
AAGCCTGACGCACCGGTAGATTCTTGGTGAGCCCGTATCATGACGGCGGCGG
GAGCTACATGGCCCCGGGTGATTTATTTTTTTTGTATCTACTTCTGACCCTTTT
CAAATATACGGTCAACTCATCTTTCCTGAGATGCGGCCTGCTTGGTATTGC
GATGTTGTCAGCTTGGCAAATTGTGGCTTTCGAAAACACAAAACGATTCTTA
GTAGCCATGCATTTTAAGATAACGGAATAGAAGAAAGAGGAAATTAAAAAA
AAAAAAAAAAACAAACATCCCGTTCATAACCCGTAGAATCGCCGCTCTTCGTG
TATCCCAGTACCAGTTTATTTTGAATAGCTCGCCCGCTGGAGAGCATCCTGAA
TGCAAGTAACAACCGTAGAGGCTGACACGGCAGGTGTTGCTAGGGAGCGTCG
TGTTCTACAAGGCCAGACGTCTTCGCGGTTGATATATATGTATGTTTGACTGC
AGGCTGCTCAGCGACGACAGTCAAGTTCGCCCTCGCTGCTTGTGCAATAATC
GCAGTGGGGAAGCCACACCGTGACTCCCATCTTTCAGTAAAGCTCTGTTGGT
GTTTATCAGCAATACACGTAATTTAAACTCGTTAGCATGGGGCTGATAGCTTA
ATTACCGTTTACCAGTGCCGCGGTTCTGCAGCTTTCCTTGGCCCGTAAAATTC
GGCGAAGCCAGCCAATCACCAGCTAGGCACCAGCTAAACCCTATAATTAGTC

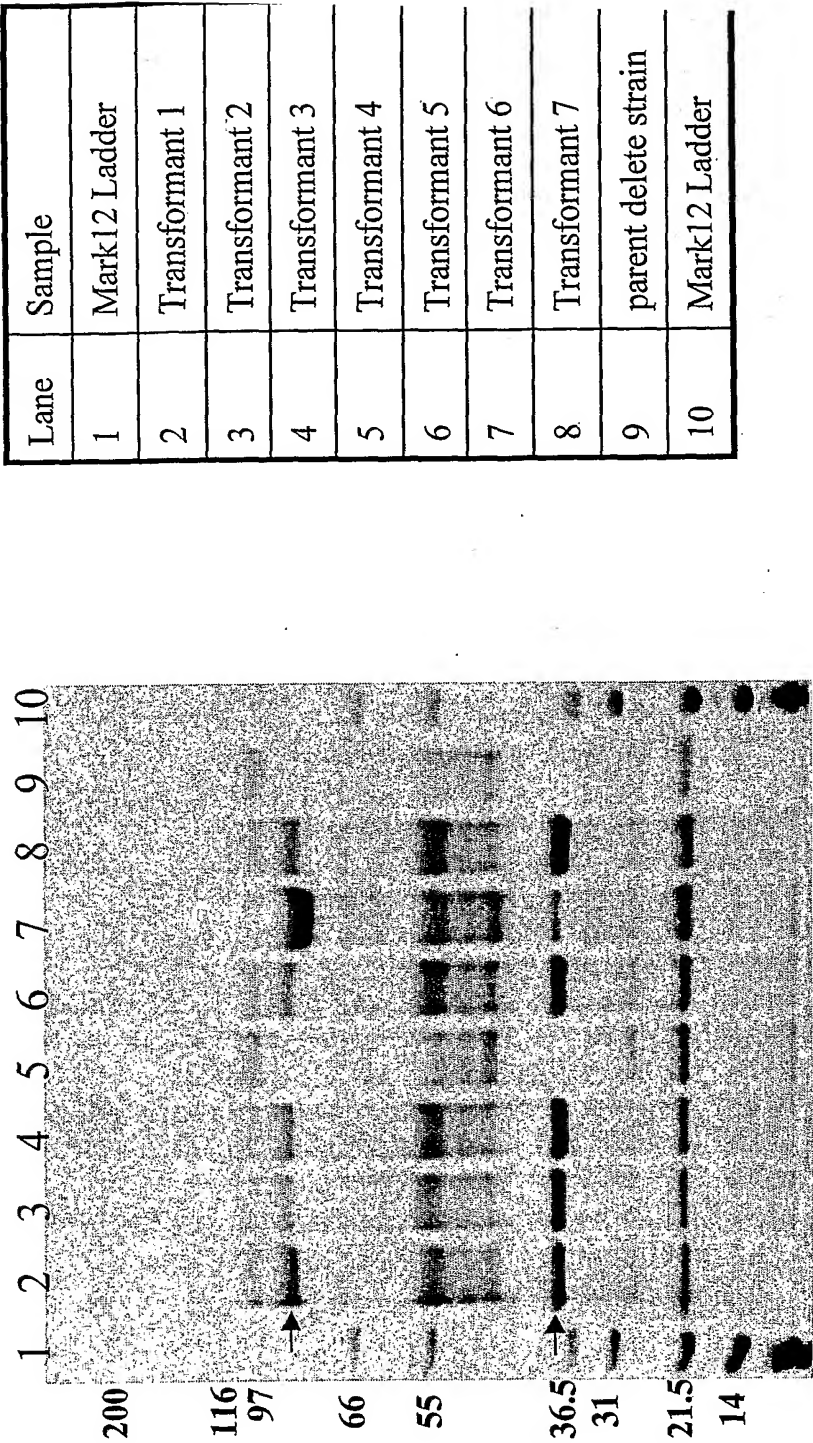
TCTTATCAACACCATCCGCTCCCCCGGGATCAATGAGGAGAATGAGGGGGAT
GCGGGGCTAAAGAAGCCTACATAACCCTCATGCCAACTCCCAGTTTACACTC
GTCGAGCCAACATCCTGACTATAAGCTAACACAGAATGCCTCAATCCTGGGA
AGAACTGGCCGCTGATAAGCGCGCCCGCCTCGCAAAAACCATCCCTGATGAA
TGGAAAGTCCAGACGCTGCCTGCGGAAGACAGCGTTATTGATTTCCCAAAGA
AATCGGGGATCCTTTTCAGAGGCCGAACCTGAAGATCACAGAGGCCTCCGCTGC
AGATCTTGTGTCCAAGCTGGCGGGCCGGAGAGTTGACCTCGGTGGAAGTTACG
CTAGCATTCTGTAAACGGGCAGCAATCGCCCAGCAGTTAGTAGGGTCCCCTC
TACCTCTCAGGGAGATGTAACAACGCCACCTTATGGGACTATCAAGCTGACG
CTGGCTTCTGTGCAGACAACTGCGCCACAGAGTTCTTCCCTGACGCCGCTCT
CGCGCAGGCAAGGGAACCTCGATGAATACTACGCAAAGCACAAGAGACCCGT
TGGTCCACTCCATGGCCTCCCCATCTCTCTCAAAGACCAGCTTCGAGTCAAGG
TACACCGTTGCCCCTAAGTCGTTAGATGTCCCTTTTTGTCAGCTAACATATGC
CACCAGGGCTACGAAACATCAATGGGCTACATCTCATGGCTAAACAAGTACG
ACGAAGGGGACTCGGTTCTGACAACCATGCTCCGCAAAGCCGGTGCCGTCTT
CTACGTCAAGACCTCTGTCCCGCAGACCCTGATGGTCTGCGAGACAGTCAAC
AACATCATCGGGCGCACCGTCAACCCACGCAACAAGAACTGGTCGTGCGGCG
GCAGTTCTGGTGGTGAGGGTGCGATCGTTGGGATTCTGTGGTGGCGTCATCGG
TGTAGGAACGGATATCGGTGGCTCGATTGAGTGCCGGCCGCGTTCAACTTC
CTGTACGGTCTAAGGCCGAGTCATGGGCGGCTGCCGTATGCAAAGATGGCGA
ACAGCATGGAGGGTCAGGAGACGGTGACAGCGTTGTCTGGGCCGATTACGCA
CTCTGTTGAGGGTGAGTCCTTCGCCTCTTCCTTCTTTTCCTGCTCTATACCAGG
CCTCCACTGTCCTCCTTTCTTGCTTTTTATACTATATACGAGACCGGCAGTCAC
TGATGAAGTATGTTAGACCTCCGCCTCTTCACCAAATCCGTCCTCGGTCAGGA
GCCATGGAAATACGACTCCAAGGTCATCCCCATGCCCTGGCGCCAGTCCGAG
TCGGACATTATTGCCTCCAAGATCAAGAACGGCGGGCTCAATATCGGCTACT
ACAACTTCGACGGCAATGTCCTTCCACACCCTCCTATCCTGCGCGGCGTGGA
ACCACCGTCGCCGCACTCGCCAAAGCCGGTCACACCGTGACCCCGTGGAACGC
CATAAAGCACGATTTTCGGCCACGATCTCATCTCCCATATCTACGCGGCTGAC
GGCAGCGCCGACGTAATGCGCGATATCAGTGCATCCGGCGAGCCGGCGATT
CAAATATCAAAGACCTACTGAACCCGAACATCAAAGCTGTTAACATGAACGA

GCTCTGGGACACGCATCTCCAGAAGTGGAATTACCAGATGGAGTACCTTGAG
AAATGGCGGGAGGCTGAAGAAAAGGCCGGGAAGGAACTGGACGCCATCATC
GCGCCGATTACGCCTACCGCTGCGGTACGGCATGACCAGTTCCGGTACTATG
GGTATGCCTCTGTGATCAACCTGCTGGATTTACGAGCGTGGTTGTTCCGGTT
ACCTTTGCGGATAAGAACATCGATAAGAAGAATGAGAGTTTCAAGGCGGTTA
GTGAGCTTGATGCCCTCGTGCAGGAAGAGTATGATCCGGAGGCGTACCATGG
GGCACCGGTTGCAGTGCAGGTTATCGGACGGAGACTCAGTGAAGAGAGGAC
GTTGGCGATTGCAGAGGAAGTGGGGAAGTTGCTGGGAAATGTGGTGACTCCA
TAGCTAATAAGTGTCAGATAGCAATTTGCACAAGAAATCAATACCAGCAACT
GTAAATAAGCGCTGAAGTGACCATGCCATGCTACGAAAGAGCAGAAAAAAA
CCTGCCGTAGAACCGAAGAGATATGACACGCTTCCATCTCTCAAAGGAAGAA
TCCCTTCAGGGTTGCGTTTCCAGTCTAGACACGTATAACGGCACAAGTGTCTC
TCACCAAATGGGTTATATCTCAAATGTGATCTAAGGATGGAAAGCCCAGAAT
CTAGGCCTATTAATATTCCGGAGTATACGTAGCCGGCTAACGTTAACAACCG
GTACCTCTAGAACTATAGCTAGCATGCGCAAATTTAAAGCGCTGATATCGAT
CGCGCGCAGATCCATATATAGGGCCCCGGGTTATAATTACCTCAGGTCGACGT
CCCATGGCCATTCGAATTCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAAT
TGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTA
AAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTC
ACTGCCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCG
GCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTC
GCTCACTGACTCGCTGCGCTCGGTCGTTTCGGCTGCGGCGAGCGGTATCAGCTC
ACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAA
AGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCG
CGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAT
CGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAG
GCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCT
TACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATA
GCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGC
TGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTA
TCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCC

ACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCT
TGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTG
CGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCG
GCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATT
ACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGT
CTGACGCTCAGTGGAACGAAAACTCACGTAAAGGGATTTTGGTCATGAGATT
ATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAAATGAAGTTTTAAAT
CAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATC
AGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTG
ACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGGCTTACCATCTGGCCCC
AGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAG
CAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTT
ATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTT
CGCCAGTTAATAGTTTTCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTG
TCACGCTCGTCGTTTGGTATGGCTTCATTACGCTCCGGTTCCCAACGATCAAG
GCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTC
CTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATG
GCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGT
GACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCG
AGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAA
CTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAG
GATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACT
GATCTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGA
AGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATA
CTCATACTCTTCCTTTTTCAATATTATTGAAGCATTATCAGGGTTATTGTCTC
ATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAATAGGGGTTC
CGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTAT
CATGACATTAACCTATAAAAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGC
GTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGT
CACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCG
TCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAAGTATGCGGCATCAGAGC

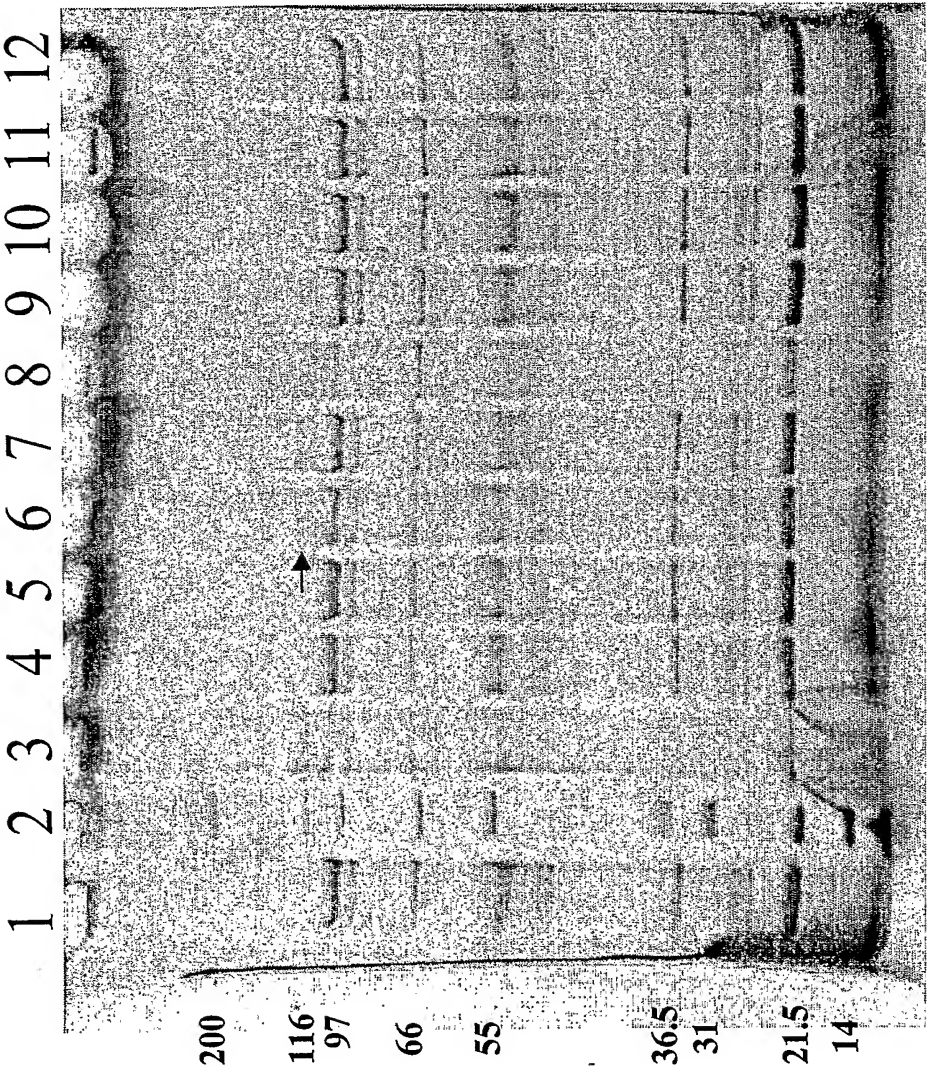
AGATTGTACTGAGAGTGCACCATAAAATTGTAAACGTTAATATTTTGTAAAA
TTCGCGTTAAATTTTTGTAAATCAGCTCATTTTTAAACCAATAGGCCGAAAT
CGGCAAAATCCCTTATAAATCAAAAGAATAGCCCGAGATAGGGTTGAGTGTT
GTTCCAGTTTGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCA
AAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC
CAAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCCT
AAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCG
AGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAG
TGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCG
CTACAGGGCGCGTACTATGGTTGCTTTGACGTATGCGGTGTGAAATACCGCA
CAGATGCGTAAGGAGAAAATACCGCATCAGGCGCCATTGCGCCATTCAGGCTG
CGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGC
TGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTT
TTCCCAGTCACGACGTTGTAAACGACGGCCAGTGCC

Figure 18
SDS-PAGE gel of supernate samples of shake flask grown *T.reesei* delete strain transformed with the fusion expression construct *cbh1-E1*



The fusion protein is indicated by the upper arrow, the cleaved E1 catalytic domain is indicated by the lower arrow

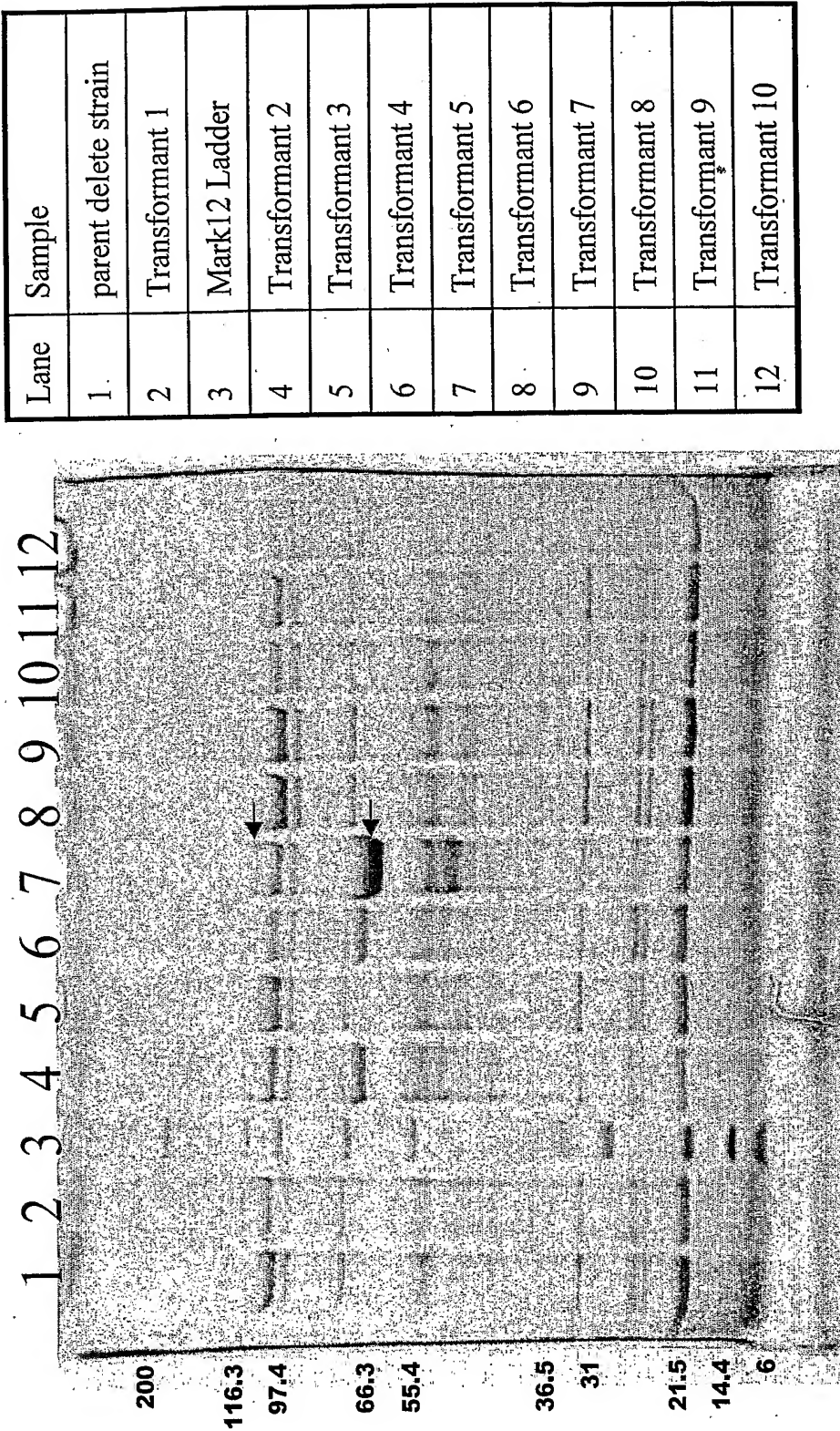
Figure 19
SDS-PAGE gel of supernate samples of shake flask grown *T.reesei* delete strain transformed with the fusion expression construct *cbh1-GH48*



Lane	Sample
1	parent delete strain
2	Mark12 Ladder
3	Transformant 1
4	Transformant 2
5	Transformant 3
6	Transformant 4
7	Transformant 5
8	Transformant 6
9	Transformant 7
10	Transformant 8
11	Transformant 9
12	Transformant 10

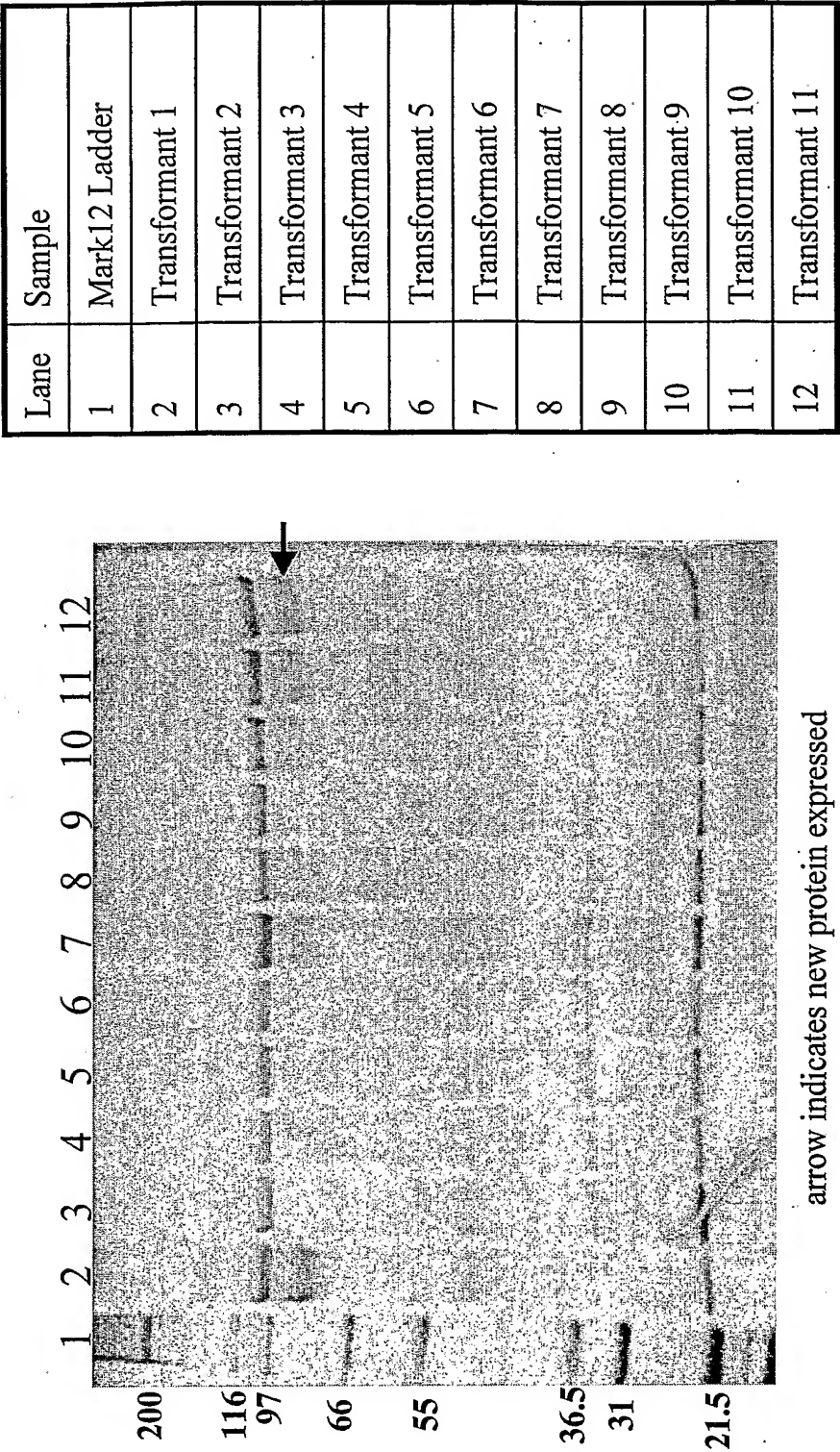
The fusion protein is indicated by the arrow

Figure 20
SDS-PAGE gel of supernate samples of shake flask grown *T.reesei* delete strain transformed with the fusion expression construct *cbh1-GH74*



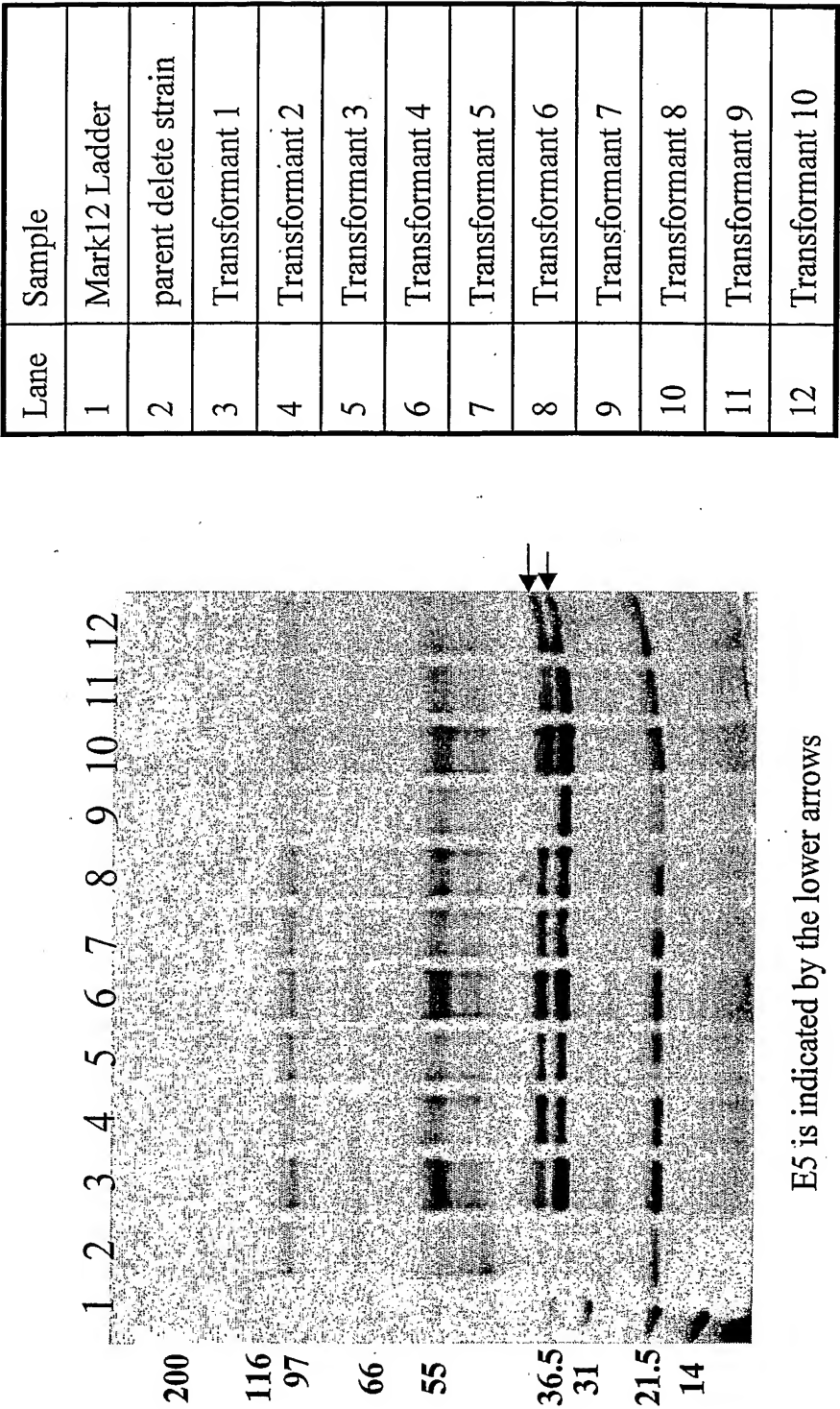
The fusion protein is indicated by the upper arrow, the cleaved GH74 catalytic domain is indicated by the lower arrow

Figure 21
SDS-PAGE gel of supernate samples of shake flask grown *T.reesei* delete strain transformed with the fusion expression construct *cbh1-E3*



arrow indicates new protein expressed
in *cbh1-E3* transformants

Figure 22
SDS-PAGE gel of supernate samples of shake flask grown *T. reesei* delete strain transformed with the fusion expression construct cbh1-E5



E5 is indicated by the lower arrows

FIGURE 23

